DOCUMENT RESUME

ED 041 325 CG 005 659

AUTHOR Chambers, David W.
TITLE Putting Down the Discovery Learning Hypothesis.
INSTITUTION American Educational Research Association, Washington, D.C.; North Carolina Univ., Chapel Hill. School of Dentistry.; North Carolina Univ., Chapel Hill. School of Education.
PUB DATE Mar 70
NOTE 24p.; Paper presented at the American Educational Research Association Convention, Minneapolis, Minnesota, March 2-6, 1970

EDRS PRICE EDRS Price MF-$0.25 HC-$1.30
DESCRIPTORS Concept Formation, *Discovery Learning, Discovery Processes, *Laboratory Experiments, *Learning, Learning Characteristics, *Learning Processes, Reinforcement, *Transfer of Training

ABSTRACT

A laboratory study was performed with four levels of discovery and two levels on an overlearning factor. An improved experimental paradigm was chosen so as to avoid a number of confusions which have obscured previous research. An operational definition was developed for discovery and guided discovery and discovery teaching, and inquiry learning. The results of the study show that overlearning has a considerably more powerful effect on transfer than does discovery. Over-learning seems to be an important condition for the transfer of a discovered principle, not because the principle is apt to be discovered during overlearning but because a certain amount of practice is necessary to make the discovered principle available for transfer. (Author)

ED041325

Putting down the discovery learning hypothesis

Paper prepared for special pre-registration session
on Learning and Cognition

Division C

American Educational Research Association Meeting

1970 Minneapolis, Minnesota

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

by

David W. Chambers

Schools of Education and Dentistry

University of North Carolina, Chapel Hill

CG 005 659

There is little comfort in the fact that studies on discovery learning are among the most popular in the literature. It is intuitively appealing that we retain and transfer material better when we learn it on our own than when we are told the same facts. But as often as not, empirical studies have failed to confirm what seems so obvious to many teachers and educational researchers.

When a flurry of studies produces no consistent results, it will no longer suffice to repeat the same studies with minor modifications. Attention should be directed towards analyzing definitions, assumptions, and methodology of studies on discovery learning. This has been the intent of recent articles by Ausubel (1961), Bruner (1961), and Wittrock (1963). The book edited by Schulman and Kieslar (1966) which is a collection of papers presented at a conference on learning by discovery is not as obviously useful. It does little more than showcase the diversity and confusion in this area of research without proceeding to the stage of positive proposals based on the discussion of participants in the conference.

The last two studies on discovery learning reported in the Journal of Educational Psychology (Guthrie, 1967; Roughead & Scandura, 1968) have focused on overcoming specific terminological or methodological problems in previous research. They both report no significant advantage in favor of the discovery group. Under these circumstances, it seem appropriate to augment the experimental evidence to be reported here with a theoretical rationale designed to justify the research design used and shed additional light on the confusion surrounding discovery learning. There are four objectives to be served in this paper: (a) provide an operational definition of discovery learning and guided discovery learning, (b) make definitional distinctions between three concepts which are often condused with discovery learning, i.e., meaningful learning, discovery teaching, and inquiry learning. (c) provide some evidence as to the educational (as opposed to statistical) significance of discovery learning, and (d) furnish additional experimental evidence which does not confirm the hypothesis that transfer is aided through discovery.

Operational definitions. It has long been known that transfer and retention properties differ with various different subject matters. The discovery learning hypothesis asks the question: Does it make any difference how something is learned? In this paper, attention will be restricted to the learning of principles or generalizations which summarize large sets of individual associations. "Grief is spelled G R I E F" and relieve is spelled R E L I E V E" are examples of individual associations which can be subsumed under the principle of "I before E, except after C." A child can spell these words correctly either because he knows a set of individual associations or a general principle. The principle can be learned in two ways -- didactically, by being told the "I before E" rule

r by discovery, by working with individual words until the general rule comes to mind.

An operational definition of discovery for principle learning can now be given. First, it will be necessary to define "learning a principle." When a person who was formerly unable to respond to new individual associations governed by an organizing generalization (as in extrapolation or application of a rule to a new situation) becomes able to correctly respond to new individual instances of the principle, we say that he has learned the principle. It will also be assumed that when this person has no means for learning the principle other than manipulation of concrete instances, he has learned the principle by discovery. In other words, if a person received didactic instruction in the use of a principle, he cannot discover this principle. (Of course, this is not true in any strict sense. As will be pointed out later, instruction is not the same thing as learning and a person may fail to understand what he is told and so he might still subsequently discover it for himself.) With this reservation in mind, it can be postulated as an operational definition that what one succeeds in learning about the organization of experience is discovered unless it is taught to him didactically.

Much school learning falls somewhere between the two pure points of didactic and discovery learning. It is, therefore, useful to extend the operational definition to cover all cases of "guided discovery" as well. On the assumption that one cannot discover what he has already been told, guidance can be identified with partial didactic instruction of the organizing principle. Explaining part of the "I before E" rule (for example, pointing out that there is a rule involving these letters) has the effect of reducing the amount left for the subject to discover about the relationship. The degree of guidance in discovery is only the proportion of the organizing principle that is didactically preempted.

Some terminological confusions. One of the reasons why there is contrary evidence in the literature on discovery learning is lack of agreement on terms. Many experimental designs have been formulated and replicated which do not provide tests of the hypothesis that it makes a difference how something is learned. In this section, the three primary sources of confusion will be identified along with a partial list of the experiments performed under research paradigms mistakenly identified with discovery learning. The three sources of confusion have been: (a) discovery learning vs. meaningful learning, (b) discovery learning vs. discovery teaching, and (c) discovery learning vs. inquiry learning.

(a) Discovery learning vs. meaningful learning. During the period of the 1930's and 1940's when educational psychologists were strongly influenced by Gestalt Psychology and Progressivism, a number of studies (Brownell & Moser, 1944; McConnell, 1934; Stacey, 1949; Stanley, 1949; Swenson, 1949; Thiele, 1939) were performed in the schools to show that curricula involving traditional rote memorization of meaningless rules

are inferior to curricula involving the discovery of meaningful organizational principles. These studies were more extensive than many to follow and to a large extent the evidence confirmed the hypotheses.

But as Ausubel points out in another context (1963), there is no necessary connection between meaningfulness and discovery on the one hand and didacticism and rote learning on the other. Meaningful and rote learning is one dichotomy which refers to the use a learner can make of material presented. Discovery and didactic learning is a second dichotomy which has already been associated with the manner in which information is learned. Although most psychologists and educators have been interested in the double dichotomy of rote didactic learning vs. meaningful discovery learning, the other two combinations are equally intelligable. For example, the theoretical structure of classical harmony can be presented in a dry but nonetheless meaningful lecture or a teenager may discover how to play the guitar "by ear."

The researchers of the 1930's and 1940's failed to observe the distinction between meaningfulness and discovery learning. In fact, their research design intentionally confounded these two variables. It would, therefore, be a very risky policy to cite these studies as evidence in favor of the alledged beneficial effects of discovery learning.

(b) Discovery learning vs. discovery teaching. This is by far the most prevalent obfuscation of the discovery learning hypothesis. The most typical experimental paradigm follows Winch (1913) who compared the effects on retention and transfer of equal exposure to a didactic and a discovery presentation format. The studies by Craig (1956), Gagné and Brown (1961), Grote (1960), Haslerud and Meyers (1958), Kersh (1958, 1962), Kittell (1957), Moss (1960), Ray (1960), Rowlett (1960), Tomlinson (1962), and Winch (1913) all make an effort to equate, in some sense, the exposure in the two treatment groups. Cronbach (1966) and Gagné (1966) urge that experimental paradigms which fail to take this precaution should be suspect.

This precaution, however justifiable for other reasons, seems to miss the point of the discovery learning hypothesis. What is at question is the manner in which material is learned and not the manner in which it is presented. My wife thinks I forgot to take out the garbage this morning and the fact that the garbage is still here strengthens her case considerably. But if she failed to tell me that she expected me to do this chore or if she told me while I was shaving in the other room, I cannot be accused of forgetting since I did not hear her in the first place. Similarly, it requires a leap of faith to assume that equal exposure to material means equal learning has taken place. Because of this flaw in the Winch paradigm, many studies should be considered tests of the relative effectiveness or speed of discovery teaching since inferences about retention and transfer are tenuous

The interactions of individual differences with mode of presentation are sufficient to discredit the assumption that "equal exposure means equal learning." To continue with instruction until even the slowest learner would have something to remember involves one in the further jeopardy of disregarding the possible interactional effects of over-learning.

We appear to be on very shakey ground indeed when we are forced to draw conclusions about what a child learned given only information on what he was exposed to. In the long run, it will be large-scale curriculum projects based on the discovery teaching paradigm which will provide the educationally significant answers for policy decisions. In the meantime, however, information gained using this paradigm should not be made to do the double duty of providing information about the manner in which people learn and the way this is reflected in what they retain.

(c) Discovery learning vs. inquiry learning. Most of the work being done today under the nominal heading of discovery learning is actually investigations into the area of inquiry learning. There would be no point in listing here the many curriculum studies which are showing with some degree of consistency that some kind of advantage does accrue to students who learn by working through exercises designed to make them discover generalizations.

What is germane, however, is to raise the question of just what that advantage might be. The discovery learning hypothesis claims that the principles which are discovered are more apt to be remembered and transferred than those which are learned didactically. But often research paradigms designed to test this hypothesis train their subjects on more than one principle and then ask that yet additional principles be used in the transfer task. This experimental arrangement leaves it open for critics to dispell results apparently favoring the discovery learning hypothesis by showing that the results could also be explained by assuming that the subject learned the generalized ability to discover principles and not merely the better use of one or two of them in particular.

The ability to discover principles, or inquiry, is an extremely useful skill that is probably best learned by practice. Dispite the fact that many results attributed to "learning by discovery" can be explained as "learning to discover" or as what Harlow calls "learning to learn" (1949, 1959), these two phenomena should not be mistakenly confused. Learning the skill of inquiry involves the learning of something which is still quite distinct from learning by means of discovery. This distinction is clarified in the context of the curriculum studies by Joseph Schwab (1966) and in the area of discovery learning by Robert Glaser's phrase "learning by discovery vs. learning for discovery" (1966). Although the curriculum reform movement based

on inquiry training may have had its origins in the laboratory studies on discovery learning, they are distinct psychological processes and it is logically fallacious to cite the impirical successes or failures in one area as evidence in the other.

An improved experimental paradigm. In order to test the hypothesis that the manner in which a principle is learned affects the transfer power of the principle it is necessary to avoid the three confusions just mentioned. The meaningfulness of the principle will not be at issue. A model must be found in which it is possible to terminate the training as soon as the subject masters the principle, but not before. For this purpose a serial anticipation routine is appropriate. Only one principle will be learned and tested for transfer in order to avoid a confusion with the skill of learning to discover.

The experimental design which most nearly meets these criteria is Judd's (1908) original dart throwing experiment. One of the problems in this old experiment was the lack of control for the psychomotor component of the task. When the subjects missed the target, it was impossible to determine whether they failed to apply the principle or applied the principle but could not throw darts well. This difficulty can be overcome by modifying the task so that subjects are only required to read the location of an object on a scale when there is an interposed lens (analogous to depth of water) and then guess where on the scale the object would appear if the distorting lens were removed. The set of individual associations to be learned in this case are the relationships between the name or number of the lenses and the kind of distortion each lens produces. These individual associations can be subsumed under an organizing principle which is directly analogous to the principle of refraction of light used in the Judd experiment.

There is one final modification on the experimental paradigm which deserves attention. Considerable press (Allen, 1969; Guba, 1969) has recently been given to the obvious fact that statistical significance is not educational significance. In the case of the discovery learning hypothesis, the question is how large must the advantage be favoring discovery before researchers make a recommendation to educators? Tests for the statistical significance of mean differences are not particularly well suited to this task. One alternative is to use a two factorial design with another independent variable that a teacher might use instead of discovery. Since the error variance would be identical for each factor, a conclusion can be reached about the relative power of discovery learning versus a plausible alternative for improving retention and transfer.

In the experiment to be reported here, overlearning was chosen as the second factor. One reason for this choice is the fact that discovery is acknowledged to be a time-consuming strategy. A prac-

tical alternative for teachers is to teach a principle didactically and then allow an opportunity for practice. But the relationship between discovery and overlearning should also be examined on theoretical grounds. Mandler (1962) summarized a number of studies using both human and subhuman subjects in which reversal transfer becomes easier with sufficient amounts of overlearning. This well-replicated result can most easily be explained by assuming that during the overlearning of the initial set of individual associations on the training task, the subjects also learned an organizational principle which transferred positively to the reversal task. Since these subjects were not instructed in the organizing principle involved (not even altered to its existence), according to the operational definition of discovery learning, they discovered the principle. There is also a considerable body of literature (Bruce, 1933; Erlebacher & Archer, 1961; Mandler, 1954a, 1954b; Mandler & Cowan 1962; Mandler & Heinemann, 1965; Mandler & Kuhlman, 1961) which shows that, for human subjects using laboratory tasks, transfer increases as overlearning does. Accordingly, it would seem reasonable to propose that overlearning is an important condition for discovery.

Hypotheses

When the proper precautions are exercised so as to equate for original mastery of the individual associations on the learning material and when the other confounding factors are controlled for, the discovery learning hypothesis can be put to a strict test. Since almost all previous studies have been confounded in some respect, there is little precedent for estimating the effects of discovery. The most parsimonious prediction would be that it makes no difference whether the organizing principle for a set of individual associations is discovered or learned didactically. A review of the literature seems to present a general picture such that as the experimantal controls are tightened the advantage claimed for the discovery group diminishes or disappears. Considering these facts, the following may be hypothesized:

> Hypothesis 1: There will be no difference in the ability to transfer a principle, regardless of whether this principle is discovered or taught didactically, or partially taught (guided discovery).

It has already been mentioned as a well-documented finding that the likelihood of positive transfer increases as a particular body of material is overlearned. Since there is no apparent reason why this experimental situation differs from those in previous studies, the following will be offered as the second hypothesis:

> Hypothesis 2: The overlearning of a set of individual associations will produce significantly greater transfer than will

simple mastery of these associations.

A third hypothesis can also be drawn from the literature on overlearning. As a possible cue to the nature of the psychological processes involved in discovery, it has been suggested that the principle used to organize a set of individual associations should be discovered during the overlearning phase of traning. This possibility can be stated in testable form as follows:

> Hypothesis 3: The overlearning of a set of individual associations will produce significantly more reports of discovered organizational principles than will simple mastery of these associations.

Subjects, instruments, and procedures

The subjects were 56 students enrolled in the Stanford University Secondary Teacher Education Program. These subjects were assigned by the draw of a card from a shuffled deck to one of eight experimental groups. The experimental procedure involved a general introductory session followed by individually administered training and transfer sessions spaced exactly one week apart. During the introductory session, the purpose of the experiment was explained, the apparatus for exposing the target against a scale with an interposed lens was demonstrated, and the experimental routine of serial anticipation was practiced. In addition, each subject was instructed in the use of a partial organizing principle determined by his assignment to a treatment group. There were four such partial organizing principles comprising the discovery factor in this experiment: The discovery group with no rules given (SP_0), a partial organizing principle relating to the direction of distortion caused by each lens (SP_1), a partial organizing principle relating to the fact that each successive pair of lenses was separated by a constant difference of two unit's distortion (SP_2), and a didactic group which recieved both partial organizing principles (SP_3).

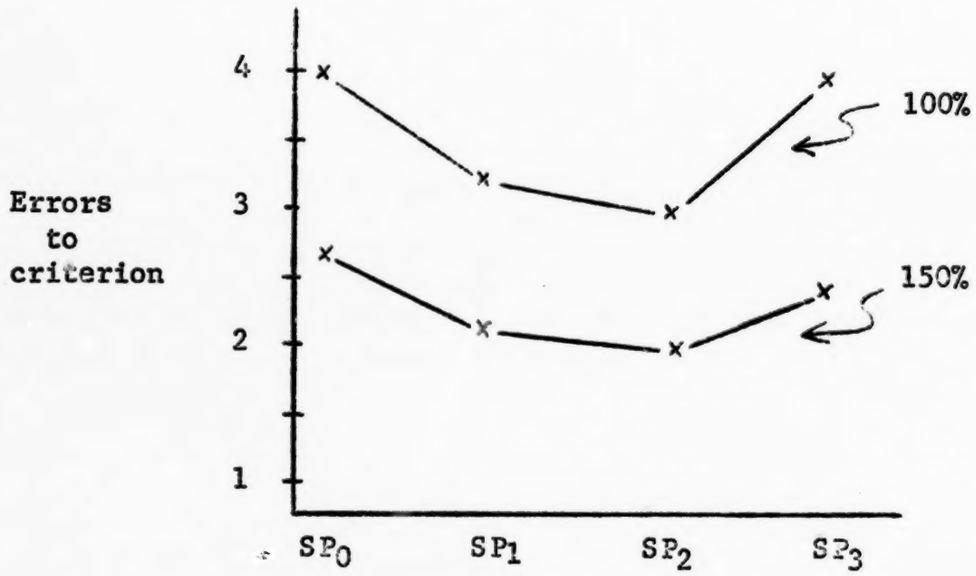
During the training session, the subject was asked to review the organizing principle he had learned and then the serial anticipation trails began. The lenses used and their distorting qualities are as follows: Third degree lens distorts three units to the right, Fourth degree lens distorts one unit to the right, Fifth degree lens distorts one unit to the left, and Sixth degree lens distorts three units to the left. If the subject had been assigned to the criterion group (100%), he would continue his training until he succeeded in making five successive correct anticipations. If his experimental assignment had been to the overlearning group (150%), he would continue the experiment until he had reached criterion and then, without interruption, he would be given an additional one trial for each two trials needed to reach criterion.

One week later, during the transfer session, the subject was told to begin the serial anticipation process again without being told the exact relationship between the task he had been trained on the week before and the transfer task. As a simple extrapolation of the principle was desired for the transfer task, a Second degree lens (five units distortion to the right) and a Seventh degree lens (five units distortion to the left) were used instead of the Fourth and Fifth degree lenses. Criterion was again considered to be five correct successive anticipations. A questionnaire was administered directly following the completion to the transfer task.

Results

The first two hypotheses can be tested directly in a two-factor analysis of variance. Figure 1 is a diagram of the mean number of errors made before reaching criterion for each of the eight treatment groups. If the discovery learning hypothesis was confirmed by these

Figure 1. Mean number of errors before reaching criterion on the transfer task for each of the eight treatment groups.

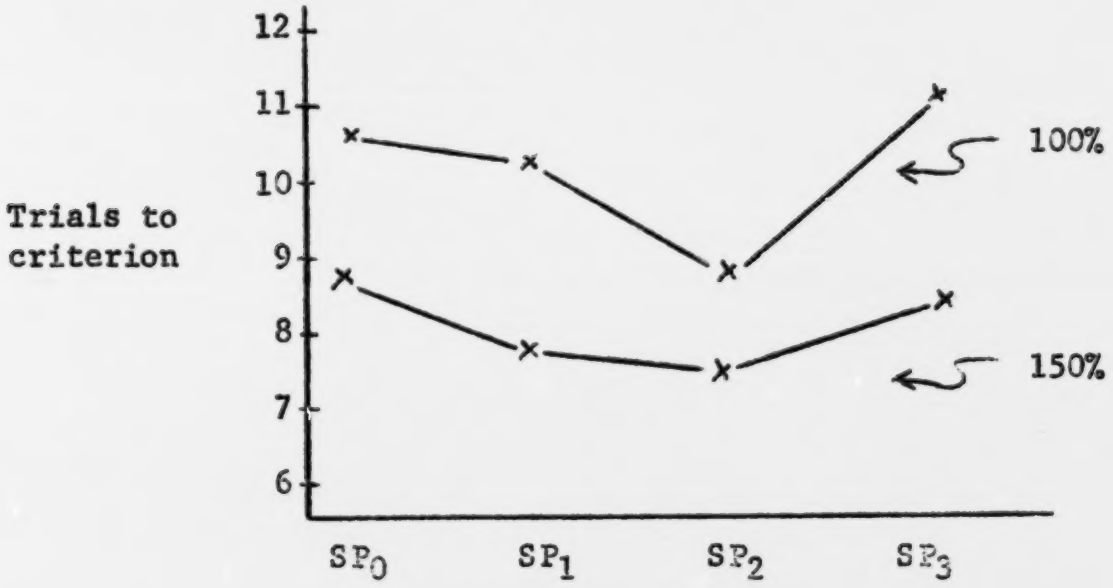


data there would be a significant increase in the number of errors from SP_0 through SP_1 and SP_2 to SP_3. This does not appear to be true of the present data and therefore the first experimental hypothesis seems to be confirmed: When groups are equated in terms of their mastery of the training task, there is no significant difference in transfer ability related to the manner in which the training task was learned. In order for the F ratio to reach significance at the α

level of 90%, a value of 2.23 is needed, The obtained F ratio is only 0.439.

The second experimental hypothesis tested in this study is that when groups are equated in terms of their mastery of the training task, there is a significant difference in transfer ability favoring those who overlearn the training task. From the diagram in Figure 1, it appears that this hypothesis is confirmed since at each level of discovery the overlearning group made fewer errors in reaching criterion. This difference is statistically significant at the $p \leq .05$ level, with an F ratio of 4.568.

Figure 2. Mean number of trails needed to reach criterion on the transfer task for each of eight treatment groups.



It is reasonable to consider an alternative to the dependent variable of number of errors made before reaching criterion on the transfer task. The number of trials required to reach criterion on the transfer task is highly correlated with the number of errors made, and Figure 2 shows that much the same relationship exists as before when the dependent variable is number of trials needed to reach criterion. The same statistical properties obtain and the hypothesis about discovery and overlearning are both confirmed. A summary of the statistical relationships for these two dependent variables and for a logarithmic transformation of the errors-to-criterion measure is found in Table 1.

With a trials-to-criterion measure, the distribution of responses is not usually normal because of ceiling effects. This is true of the dependent variable in this study which has a distribution skewed

toward few trials. A logarithmic transformation is often the most orderly description for such skewed data. Therefore, the same tests were recalculated for the double transformation of adding one to each error-to-criterion score and then taking the logarithm to the base 10 of this new score. As can be seen in Table 1, the same statistical relationships hold and again the discovery learning hypothesis fails to find confirmation while the overlearning effect is pronounced.

Table 1. Summary of the analysis of variance tables for the two-factor analysis of variance tests of Hypothesis 1 and Hypothesis 2 with three different dependent variables.

Table 1a. Analysis of variance table for number of errors before reaching criterion.

Source	df	SS	MS	F
Overlearning	1	24.446	24.446	4.568*
Discovery	3	7.054	2.351	0.439
Interaction	3	0.624	0.208	0.039
Within cells	48	256.856	5.351	
Total	55	288.981		

Table 1b. Analysis of variance table for number of trails needed to reach criterion.

Source	df	SS	MS	F
Overlearning	1	66.445	66.445	4.279*
Discovery	3	22.339	7.446	0.481
Interaction	3	4.487	1.496	0.099
Within cell	48	722.852	15.059	
Total	55	816.125		

Table 1c. Analysis of variance table for transformed number of errors before reaching criterion (transformed by adding a constant of one to each score and then taking the logarithm to the base 10 of the new score).

Source	df	SS	MS	F
Overlearning	1	0.590	0.590	5.246*
Discovery	3	0.387	0.129	1.146
Interaction	3	0.005	0.002	0.015
Within cell	48	5.401	0.113	
Total	55	6.383		

* = significant at the $p \leq .05$ level

It is impossible to make a direct test of the third experimental hypothesis from dependent variables such as errors or trials to criterion. For this purpose it is necessary to turn to the questionnaire data. The relevant questions are the following: "What regularities, if any, did you notice in the first experiment a week ago?" "When did you recognize there regularities?" and "Did these regularities seem to reoccur in the experiment today?"

The three tables below show how each question was answered by subjects in each of the six appropriate treatment groups. By referring to the codes printed beside each table, three conclusions can be reached: In the first place, it is possible to make a statistical test of the hypothesis relating overlearning to the reported discovery of organizing principles (Hypothesis 3). In Table 2a the regularities which subjects reported discovering during the training session are listed for each of the six treatment groups. The code numbers of six or higher (including zero) indicate reported regularities such as "half of the distortions were to the right and half to the left (code score = 6) or the distortions were in a mirror image configuration (code score = 8)." Coded scores of six or higher would be helpful in making a transfer to the second, transfer task. In order to substantiate the third hypothesis it is necessary to show that subjects in the overlearning condition report more regularities with a code score of six or higher than the subjects in the criterion condition do. An inspection of Table 2a shows that exactly nine subjects in each treatment group recieved such a score.

It might, however, be suggested as a more powerful formulation of the third hypothesis that the overlearning subjects learned "more useful" rules. By assuming that the code six rule is less useful than the code seven rule, and so forth for each of the coded rules, it is possible to test the third hypothesis in its more powerful form throught the use of the Mann-Whittney U test for differences in two independent samples. For the conditions found in this experiment, a U statistic of less than or equal to 21 would be needed to reject the hypothesis that there is no difference in the usefulness of the reported regularities at the $p \leq .05$ level. The computed U for this experiment is 37.5. This means that the data do not confirm the third hypothesis in this case; overlearning of a set of individual associations did not produce significantly more reports of discovered organizing principles (or reports of "more useful" organizing principles) than did simple mastery of those associations.

Secondly, the time of discovery of these principles is also comparable as Table 2b shows. The code numbers for the overlearning group are generally smaller because the time of discovery is recorded relative to the overall number of trails given. This difference reflects an advantantage for the overlearning group since they could practice with the principles they had discovered while almost half of the criterion group subjects discovered their organizing regularities "late in or after the training session."

Table 2. Questionnaire responses to questions pertaining to the discovery of regularities during the training session.

Table 2a. Reported regularities discovered during the training task.

	SP_0	SP_1	SP_2
100%	1 8 2 8 6 9 8	1 8 3 8 3 8 8	1 3 1 3 1 4 1
150%	2 9 5 9 5 0 8	1 7 3 8 5 8 5	1 4 1 6 1 8 1

Code

1. no regularities
2. cannot remember
3. irrelevant regularities
4. same rule as that taught
5. list of distortions
6. groupings of + and -
7. SP_2
8. mirror image distrotions
9. SP_1 plus two values or mirror images
0. SP_2 plus + and - groups or mirror image

Table 2b. Time at which the regularity was discovered.

	SP_0	SP_1	SP_2
100%	1 8 1 5 3 5 4	1 2 1 3 1 4 2	1 1 1 1 1 1 1
150%	2 3 3 4 3 5 3	1 3 1 4 2 4 2	1 1 1 2 1 2 1

Code

1. no regularities, cannot remember, irrelevant, or same as principle
2. during first 4 anticipations
3. during second 4
4. half-way through training task
5. late in or after the training task

Table 2c. Recognition that the discovered regularities reoccured in the transfer task.

	SP_0	SP_1	SP_2
100%	1 6 5 6 6 6 6	1 6 1 6 1 6 6	1 1 1 1 1 1 1
150%	3 6 3 6 3 6 4	1 6 1 6 6 6 6	1 1 1 6 1 6 1

Code

1. no regularities, cannot remember, irrelevant, or same as principle
2. no recognition
3. partial recognition
4. probable recognition
5. different regularity reported
6. recognition

Table 3. Questionnarie responses to questions pertaining to the initial intelligibility of the partial organizing principles and their helpfulness during the training task.

Table 3a. Apparent intelligibility and agreement of the partial organizing principles.

	SP_1	SP_2	SP_3	Agreement (SP_3 only)
100%	1 1 1 2 1 3 1	1 1 1 1 1 3 1	1 1 1 1 1 3 1	1 2 1 2 1 3 1
150%	1 1 1 1 1 1 1	1 1 1 2 1 2 1	1 1 1 1 1 1 1	1 1 1 1 1 2 1

Code
1. yes
2. unclassified
3. no

Table 3b. Reported helpfulness of the SP_1 principle.

	SP_1	SP_3
100%	1 1 1 2 1 4 1	1 2 1 3 1 7 1
150%	1 1 1 3 1 3 1	1 2 1 2 1 7 1

Code
1. helped -- I saw at the beginning that it was applicable
2. helped -- applicability not recognized immediately
3. helped -- but confusing at first
4. neither helped nor hurt
5. hurt -- seemed irrelevant
6. hurt -- was confusing
7. some other relationship

Table 3c. Reported helpfulness of the SP_2 principle.

	SP_2	SP_3
100%	1 2 2 3 2 4 2	1 2 1 2 2 7 2
150%	1 3 1 4 1 7 2	1 2 1 7 1 7 1

Code
1. helped -- I saw at the beginning that it was applicable
2. helped -- applicability not recognized immediately
3. helped -- but confusing at first
4. neither helped nor hurt
5. hurt -- deemed irrelevant
6. hurt -- was confusing
7. some other relationship

Finally, understanding the relation above helps to explain the fact that when the subjects were asked if they recognized that the principle whey had discovered during the training session was applicable on the transfer task, over 50% more criterion than overlearning subjects said they did not recoginze the applicability of what they had discovered. This relationship is reflected in an almost identical proportion of the subjects who completely filled the answer to the first question on the questionnaire (by inference, the recognition of a common organization in both the training and the transfer task is likely then the distortions are completely listed for the six lenses involved.) Unfortunately, the differences in both cases are not large enough to detect with a χ^2 statistic. But the fact that the overlearning group actually did perform better on the transfer task also must be considered to add its weight to the speculation that the overlearning group more often recognized the applicability of what they had learned because they had more opportunity to practice what they had discovered.

Two questions were also asked about the helpfulness of the partial organizing principles offered during the instruction session. These two questions were not asked of the SP_0 group. "In the group instruction session before the first experiment, you were taught a rule for these experiments. Did this rule seem to make sense to you at the time when you first learned it?" and a question asking the subject to check the appropriate alternative from among several descriptions of how this principle might have been helpful or confusing during the training session. In addition to having a separate set of alternative for each rule, the SP_3 questionnaire also contained a question pertaining to the perceived consistency of the two principles during the instruction session. The responses to these questions are shown in Table 3.

There are no unexpected results in these tables. The instructions on the principles seem to have been well understood and thought helpful. It is also interesting to note in Table 4 that there are no consistant and significant relationships between the reported helpfulness of a rule and the number of errors made before reaching criterion. Perhaps this is a useless question to ask since no real standard is provided for "helpfulness." Perhaps no differences are found because no one reported that the rules were not helpful in some form or another. What is, however, of some interest is the relationship between the number of errors to criterion and reports of whether or not the rules were understood (see Table 4c). When the rule was said to have made sense the average number of errors was 2.72, when the rule was said not to have made sense the errors rose to 4.00, and when the response could not be classified an intermediate number of errors (3.33) occured. It is also interesting to note that of the six cases where the subject could not report that the rule was understood, the number of errors before reaching criterion was nearly twice as large in the criterion groups as in the discovery groups. No conclusion

should be drawn from an isolated fact about six subjects, however, this result is consistent with the advantage ascribed to the practice afforded the overlearning groups with respect to their ability to recognize the applicability of the principles they had discovered.

Table 4. The relationship between the reported intelligibility and helpfulness of the partial organizing principles and their transfer power, measured in terms of errors made before reaching criterion.

Table 4a. Reported intelligibility of the partial organizing principles and errors made before reaching criterion.

Reported intelligibility *	1	2	3
Number of cases	36	3	3
Average errors	2.72	3.33	4.00

Table 4b. Reported helpfulness of the partial organizing principle SP_1 and errors made before reaching criterion.

Reported intelligibility	1	2	3	4	5	6	7
Number of cases	18	4	3	1			2
Average errors	3.16	2.50	3.67	4.00			2.00

Table 4c. Reported helpfulness of the partial organizing principle SP_2 and errors made before reaching criterion.

Reported intelligibility	1	2	3	4	5	6	7
Number of cases	10	10	2	2			4
Average errors	2.40	3.70	3.50	2.50			2.00

* codes are listed in Table 3.

Two final question of the sort discussed above will point to a similar conclusion. All groups, except for the SP_0 subjects, were asked: "Did you realize that the experiment today also involved the rules you learned in the instruction session?" and "When did you realize this?" The SP_3 questions were asked separately for each principle. The responses to these questions and the average across all subjects for errors to criterion are listed by category in Table 5.

Table 5. Questionnaire responses to questions pertaining to the recognition that the learned partial organizing principles were applicable in the transfer task.

Table 5a. Reported recognition of applicability.

			(for SP_3 S's)		
	SP_1	SP_2	SP_1	SP_2	Codes
100%	1 1 1 1 1 1 1	1 1 1 1 1 3 1	1 1 1 1 1 1 1	1 1 1 1 1 1 1	1. yes 2. unclassifiable 3. no
150%	1 1 1 1 1 2 1	1 1 1 1 1 2 1	1 1 1 1 1 1 1	1 1 1 1 1 1 1	

Table 5b. Time of reported recognition.

			(for SP_3 S's)		
	SP_1	SP_2	SP_1	SP_2	Code
100%	2 8 2 8 4 8 8	2 2 2 2 2 4 2	2 2 2 6 2 8 2	2 3 2 4 2 6 3	1. no recognition 2. during first 4 anticipations 3. during second 4 4. half-way 6. after transfer task 7. on questionnaire 8. guessed prior to transfer task
150%	1 2 2 8 2 8 2	1 2 2 2 2 3 2	2 2 2 2 2 4 2	2 2 2 7 2 7 2	

Table 5c. Relationship between recognized applicability of the partial organizing principle and errors before reaching criterion.

Recognition	1	2 or 3
Number of cases	53	3
Average errors	3.19	4.00

Table 5d. Relationship between time of recognition and errors made before reaching criterion.

Time of recognition	1	2	3	4	6	7	8
Number of cases	2	36	3	4	3	2	7
Average errors	1.50	2.40	4.70	6.75	3.00	6.00	2.50

An inspection of Table 5a will reveal that subjects in the overlearning and criterion treatment groups do not differ from each other with respect to frequency with which they failed to recognize the applicability of the partial organizing principles they had been taught. Only three subjects in the experiment reported failing to make this recognition.

It should not be concluded, however, that presence or absence and time of recognition of applicability are unrelated to transfer score. In Table 5c it can be seen that the subjects who could not be certain that they recognized the applicability of the partial organizing principles they had been taught made more errors on the transfer task before they reached criterion. Time of recognition also affects transfer power, as can be seen in Table 5d. Although the two subjects who reported no recognition averaged only one and a half errors before reaching criterion, the other error scores are what might be predicted; subjects who recognized applicability at the beginning of or prior to the transfer task had error scores of 2.40, 4.70, and 2.50. Subjects who made this recognition late in or after the transfer task or while filling out the questionnaire averaged 6.75, 3.00, and 6.00 errors before reaching criterion.

Little confidence should be placed in these data because the orregular distribution of subjects leaves some categories with very few subjects in them. It may be concluded, in the form of a hypothesis for future verification, that presence or absence and time of recognized applicability of a previously learned organizing principle affects transfer power. The effects of overlearning on such recognition cannot be assessed from the present data because of the ceiling effect. (The fact that almost all subjects recognized the applicability of the partial organizing principles provides a conformation of the uniform effectiveness of the intruction sessions.) In addition, the fact that only three subjects reported that they failed to recognize the applicability of the partial organizing principle <u>while</u> they were performing the transfer task casts valuable light on the problem of meaningful vs. rote learning of individual associations. Although it remains a possibility that these subjects performed the transfer task using individual associations while tacitly recognizing the applicability of a useful organizing principle, it would appear more likely that the partial organizing principles were learned meaningfully.

Interpretation of results

The results of this experiment provide evidence which is consistant with the first two of the three hypotheses stated. Using either trials or errors to criterion on a near transfer task as the dependent variable shows that the overlearning factor achieved statistical sig-

nificance at the $p \leq .05$ level while the discovery factor failed to do so. On the whole the critical ratio for the overlearning effect was from five to ten times as large as the critical ratio for degree of discovery. At the very least, it must be conceded that the discovery learning effect, if it exists at all, is not nearly as strong as is the effect of overlearning. If transfer is to be the measure of knowledge, "practice beyond the point of initial mastery" seems to be a better rule to recommend to teachers than "let the student discover the generalization for himself."

The third hypothesis (relating reported discoveries to overlearning) was not confirmed, yet an important insight was gained into the processes involved in transfer based on the use of an organizing principle. There were several lines of evidence leading from the questionnaire responses to the supposition that practice with an organizing princip;e improves the transfer power of that principle. There were no differences between overlearning and criterion subjects with respect to reported regularities being disoovered or in the number of trials on the training task that preceeded such discoveries. However, the overlearning subjects did have the benefit of practice with the regularities they had discovered, This increased opportunity for practice was correlated with reports of recognizing the applicability of the discovered principle on the transfer task, with an integration of the individual associations for both experimental sessions in response to the questionnaire, and with a low number of errors made before reaching criterion on the taansfer task. These data also support the belief that meaningful principles rather than sets of individual associations are at the foundation of such transfer. When subjects reported that the instruction they received on the partial organizing principles was unclear, those who had the advantage of additional practice performed better on the transfer task. Finally, by analyzing the responses to questions about realizing that the given partial organizing princip;es were applicable in the transfer task, it can be seen that transfer scores improve when the applicability is recognized.

Of course, it must be cautioned that no statistically conclusive test of the effect of practice and recogized applicability was made. But the convergence of several lines of evidence should be taken to justify entertaining a new hypothesis about the relationship between practice and recognized applicability as an intervening variable which sheds some light on organizing principles. This hypothesis would have several direct connections with the literature on transfer. For example, the reversal transfer effects reported by Mandler (1962) might be analyzed in these terms. There is also a connection here with Asch's (1969) recent theoretical paper on the problem of associations. In this article, Asch summarizes, "Of main consequence is the finding that there was no recall without recognition. When recognition failed, subjects had to relearn a well-formed and retained association as if it

were completely new" (Asch, 1969, p. 101). Finally, Judd's early admonition bears repeating: "Note that theory was not of value until it was backed by practice, but when practice and theory were both present the best adjustment was rapidly worked out" (Judd, 1908. p. 37). This seems also to justify the fact that, in the present experiment, there was no relationship between didactic instruction on organizing principles and speed of mastering the training task.

Summary

A laboratory study was performed with four levels of discovery and two levels on an overlearning factor. An improved experimental paradigm was chosen so as to avoid a number of confusions which have obscured previous research. An operational definition was developed for discovery and guided discovery and discovery learning was distinguished from meaningful learning, discovery teaching, and inquity learning. The results of the study show that overlearning has a considerably more powerful effect on transfer than does discovery. Overlearning seems to be an important condition for the transfer of a discovered principle -- not because the principle is apt to be discovered during overlearning but because a certain amount of practice is necessary to make the discovered principle available for transfer.

References Cited

Allen, D. On the social insignificance of statistical significance. Educational Researcher, 1969, 20, 3-4.

Asch, S. E. A reformulation of the problem of associations. American Psychologist, 1969, 24, 92-102.

Ausubel, D. P. learning by discovery: Rationale and mystique. Bulletin of the National Association of Secondary School Principles, 1961, 45, 18-58.

Ausubel, D. P. The psychology of meaningful verbal learning: An introduction to school learning. New York: Grune & Stratton, 1963.

Brownell, W. A., & Moser, A. G. Meaningful versus mechnaical learning: A study of grade three subtraction. Duke University Research Studies in Education, 1949, No. 8.

Bruce, R. W. Conditions of transfer of training. Journal of Experimental Psychology, 1933, 16, 343-361.

Bruner, J. S. The act of discovery. Harvard Educational Review, 1961, 31, 21-32.

Craig, R. C. The transfer value of guided learning. New York: Columbia University Bureau of Publications, 1953.

Craig, R. C. Directed versus independent discovery of established relationships. Journal of Educational Psychology, 1956, 47, 223-234.

Cronbach, L. J. The logic of experiments on discovery. In L. S. Schulman & E. R. Keislar (Eds.), Learning by discovery: A critical appraisal. Chicago: Rand McNally, 1966, pp. 77-92.

Erlebacher, A., & Archer, E. J. Preserveration as a function of degree of learning and percent of reinforcement in card sorting. Journal of Experimental Psychology, 1961, 62, 510-517.

Gagne, R. M. Varieties of learning and the concept of discovery. In L. S. Schulman & E. R. Keislar (Eds.), Learning by discovery: A critical appraisal. Chicago: Rand McNally, 1966, pp. 135-150.

Gagne, R. M., & Brown L. T. Some factors in the programming of conceptual learning. Journal of Experimental Psychology, 1961, 62, 313-321.

Glaser, R. Variables in discovery learning. In L. S. Schulman & E. R. Keislar (Eds.), Learning by discovery: A critical appraisal. Chicago: Rand McNally, 1966. pp. 13-26.

Grote, C. N. A comparison of the relative effectiveness of directed-detailed and directed discovery methods of teaching selected principles of mechanics in the area of physics. Unpublished doctoral dissertation, University of Illinois, 1960.

Guba, E. G. Significant differences. Educational Researcher, 1969, 20, 4-5.

Guthrie, J. T. Expository instruction versus a discovery method. Journal of Educational Psychology, 1967, 58, 45-49.

Harlow, H. F. The formation of learning sets. Psychological Review, 1949, 56, 51-65.

Harlow, H. F. Learning set and error factor theory. In S. Koch (Ed.), Psychology: A study of a science. Vol. 2. New York: McGraw-Hill, 1959. pp. 492-537.

Haslerud, G. M., & Meyers, S. The transfer value of given and individually derived principles. Journal of Educational Psychology, 1958, 49, 293-298.

Judd, C. H. The relation of special training to general intelligence. Educational Review, 1908, 36, 28-42.

Kersh, B. Y. The adequacy of meaning as an explaination for the superiority of learning by independent discovery. Journal of Educational Psychology, 1958, 49, 282-292.

Kersh, B. Y. The motivating effect of learning by discovery. Journal of Educational Psychology, 1962, 53, 65-71.

Kittell, J. E. An experimental study of the effect of external direction during learning on transfer and retention of principles. Journal of Educational Psychology, 1957, 48, 391-405.

Mandler, G. Response factors in human learning. Psychological Review, 1954, 61, 235-244. (a)

Mandler, G. Transfer of training as a function of response overlearning. Journal of Experimental Psychology, 1954, 47, 411-417. (b)

Mandler, G. From association to structure. Psychological Review, 1962, 69, 415-427.

Mandler, G., & Cowan, P. A. Learning of simple structures. Journal of Experimental Psychology, 1962, 64, 177-183.

Mandler, G., & Heinemann, S. H. Effect of overlearning of a verbal response on transfer of training. Journal of Experimental Psychology, 1956, 52, 39-46.

Mandler, G., & Kuhlman, C. K. Proactive and retroactive effects of overlearning. Journal of Experimental Psychology, 1961, 61, 76-81.

McConnell, T. R. Discovery vs. Authoritative identification in the learning of children. University of Iowa Studies in Education, 1934, 9, 13-26.

Moss, J. An experimental study of the relative effectiveness of the direct- detailed and the directed discovery methods of teaching letterpress imposition. Unpublished doctoral dissertation, University of Illinois, 1960.

Ray, W. E. An experimental comparison of direct-detailed and direct discovery methods of teaching micrometer principles and skills. Unpublished doctoral dissertation, University of Illinois, 1957.

Roughead, W. G., & Scandura J. M. What is learned in mathematical discovery? Journal of Educational Psychology, 1968, 59, 283-289.

Rowlett, J. D. An experimental comparison of direct-detailed and directed discovery methods of traching orthographic projection principles and skills. Unpublished doctoral dissertation, University of Illinois, 1960.

Schwab, J. J. The teaching of science as enquiry. In J. J. Schwab & P. F. Brandwein, The teaching of science. Cambridge: Harvard University Press, 1966. pp. 1-103.

Stacey, C. L. The law of effect in retained situations with meaningful material. In E. J. Swenson (Ed.), Learning theory in school situations. Minneapolis: University of Minnesota Press, 1949. pp. 74-103.

Stanley, J. C. The role of instruction, discovery, and revision in early learning. Elementary School Journal, 1949, 40, 455-458.

Swenson, E. J. Organization and generalization as factors in learning, transfer, and retroactive interference. In E. J. Swenson (Ed.), Learning theory in school situations, Minneapolis: University of Minnesota Press, 1949. pp. 9-39.

Thiele, C. L. The contribution of generalization to the learning of addition facts. Teachers College Contributions to Education, 1939, No. 763.

Tomlinson, R. M. A comparison of four presentation methods for teaching complex technical material. Unpublished doctoral dissertation, University of Illinois, 1962.

Winch W. H. Inductive versus deductive methods of teaching: An experimental research. Baltimore: Warwick & York, 1913.

Wittrock, M. C. Verbal stimuli in concept formation: Learning by discovery. Journal of Educational Psychology, 1963, 54, 183-190.